Helping the Invisible Hand
A Theory of Endogenous Mutual Concern

Martin Kolmar*
University of Konstanz

November 15, 2001

Abstract

This paper develops a theory of endogenous mutual concern. The mutual concern or the morality of economic agents is seen as a means to reduce the inefficiencies of markets in an environment where the enforcement of property rights is costly. We show that rational agents have an incentive to invest resources to create a mutual concern. Furthermore, the strength of the mutual concern depends on the technologies of production and appropriation and defense. People in economies with a high natural productivity invest more resources in the creation of a mutual concern than people in economies with low natural productivity. By the same token, morality is a means to reduce the frictions in an economy only if the decisiveness of appropriation and defense is relatively similar.

Keywords: Productivity, Anarchy, Morality
JEL classification: D74, K42, Z13

---

*I thank Friedrich Breyer and Herschel Grossman for helpful comments. The kind hospitality of Brown University is gratefully acknowledged. Research on this paper has been supported by the German Research Foundation. Send mail to: University of Konstanz, Department of Economics, Box D 136, D-78457 Konstanz, email: Martin.Kolmar@uni-konstanz.de

1 Introduction

„It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest.“

(Adam Smith, The Wealth of Nations, p. 14).

Adam Smith's powerful insight that even mutually unconcerned individuals can be guided to their mutual advantage by the invisible hand of the market is one of the most important, influential, and most reliable results in economics. It would therefore be bold to challenge this view. The purpose of this paper is rather to add a new perspective. It is argued that initially unconcerned economic agents have an incentive to invest in the creation of a mutual concern if the enforcement of property rights is costly. In a nutshell: in an economy with transaction costs the invisible hand has difficulties implementing the efficient prices, the mutual concern of the agents helps it implementing them.

This paper is not the first to qualify Adam Smith's insight. There are different ways to ensure that markets work. Market transactions can rely on the official law-enforcement system, reputation, or the intrinsic motivation of the trading partners to fulfill one's obligations. A lot of potential market transactions cannot be supported by official law enforcement because the use of the courts is too expensive or the information necessary is unverifiable. By the same token, reputation depends on repeated interaction, which is not always given. In case that the interaction is not repeated a market cannot emerge in the absence of the right morality.

Empirically we observe market transactions in situations where economic theory that focuses on egoistic motivations would predict that markets cannot emerge because neither law-enforcement nor reputation mechanisms work. This is especially true for trades of small value and with serious verification problems. A stranger in a city buying some food in a grocery store has basically no way to proof the amount of money he gave to the dealer. Nevertheless, most of the time the dealer returns the change. This is an example of a market transaction that is neither supported by the rule of law nor by repeated interaction. A number of everyday transactions are of this character because even without verification problems the opportunity costs

of going to the courts may be prohibitive. The findings of experimental economics, especially on the ultimatum game further stress the point that markets can be successful ins environments where classical economic theory would predict that market transactions must fail.[1]

The institutional alternative to the public enforcement of law is private enforcement by defensive investments against the appropriation by others. The stranger in the example could carry a gun in order to make his desire more credible to get back his change. From the point of view of the customer, carrying a gun may make the idea plausible that the food has been overpriced from the beginning and that he actually wants to get back a little more than his change. The dealer can defend against the appropriation if he himself invests in defense. Hence, the basic structure of self-enforced property rights is that of a contest.[2] As it has long been documented in the literature, contests are inefficient because part of the potential welfare is wasted by the attempt to possess it.[3] This loss of welfare puts a limit to the extend to which markets can extend by means of self enforcement.

In this paper we analyze the role of a morality as a means to reduce the resource costs of trade. The idea that morality can be a solution to a cooperation problem has obtained a considerable amount of attention in the literature on political philosophy. The idea has been most forcefully proposed by Gauthier (1986), other advocates of the idea are Narveson (1988) and Danielson (1996). If the process of creating a disposition to respect socially agreeable strategies is successful, the cooperation problem is solved without any recourse to outside enforcement.

This idea is in a major respect incomplete and misleading. It is important to note that the creation of a common morality is itself a process that requires the investment of scarce resources. The economic costs of the creation of social norms are generally ignored by the literature on political philosophy. If a moral disposition could be changed without any costs, rational agents would have no interest not to perfectly align their preferences. This would eliminate every potential for conflict in

[1] For a survey on the literature on experimental economics see Kagel and Roth (1997).

[2] See Grossman 2001.

[3] For example Skaperdas 1992 among others.

a society. The crucial question therefore becomes why this is not done. The key to understand this question is that the creation of morality is itself a costly process. The resource costs of creating a morality strong enough to overcome a cooperation problem might use so much resources that the society ends up worse off. We therefore have to understand the costs of creating a morality compared to the costs of other institutional alternatives in order to understand its economic role.

In order to do so we develop a model where economic agents can legally buy or illegally appropriate goods and defend against appropriation. The conditions are analyzed under which rational agents confronted with the waste of resources have an incentive to invest resources in the creation of a mutual concern. Mutual concern is defined in the sense that agents abstain from appropriating goods even if no law-enforcement or reputation mechanism constrains their behavior.

The approach takes as given that the agents have the potential for a moral sentiment (moral disposition), but that the extend to which this moral sentiment influences the behavior of the agents is endogenous. It closes a gap in the literature on the emergence of cooperation. The existing literature can be divided in two different categories. The first strand of the literature assumes that the behavioral influence of altruism or other factors driving cooperation are given and exogenous for the individuals. The two distinctive features of this class of models are (a) that cooperation is explained by the motivation of the agents and (b) that the altruistic or moral motivation is exogenous from the point of view of a single agent. Hamilton's (1978) famous kin-selection theory is an example for this literature: let B be the benefit for agent 2 and C the costs for agent 1 resulting from an action undertaken by agent 1. Agent 1 will act if $C < rB$, where r measures the genetic relatedness between both agents[4] (Hamilton's Rule, Hamilton 1964). What is important for our purpose is that the measure of relatedness r is exogenous for the individuals. Evolutionary theories of cooperation have a similar structure, however, the focus of these theories in on the explanation why und under what conditions altruistic agents are evolutionary stable. In these models altruism is endogenous as an evolutionary pattern but exogenous from the point of view of a single individual or a single

[4] For example for siblings it is equal to 1/2.

generation.

The other strand of the literature focuses on repeated interaction as a clue to understand the emergence of cooperation.[5] These are theories of cooperation and not theories of altruism and morality because every agent's motivation is egoistic. Hence, cooperation emerges as a rule of action. The indefinitely repeated interaction enables the agents to formulate contingent strategies, which help solving a prisoner's dilemma or other types of cooperation problems.[6] One focus of this literature is on the cognitive prerequisites for direct or indirect reciprocity as strategies to overcome the cooperation failure.[7]

This paper has an intermediate position, both with respect to the methodological approach as well as with respect to the phenomena that are to be explained by the theory. Individualistic theories can explain certain phenomena either as a consequence of the motivation of the agents or as a consequence of the restriction the agent faces. The first class of models focuses on the motivation or preferences of the agents, whereas the second class of models focuses on the rules or restrictions of the agents as the explanatory variable of empirical patterns. The dividing line between restriction and motivation is to some extend arbitrary because all we can observe is behavior and a lot of restrictions are impossible to measure, especially social norms. In the case of social norms it is impossible to classify behavior as a result of restrictions or as a result of motivation as long as nobody breaks the norm because potential punishment are nowhere codified. For the purpose of this paper the following definition is used. Cooperative behavior that rests on any explicit or implicit mechanism of external punishment is a result of restrictions, whereas cooperative behavior that exists even in the absence of an explicit or implicit mechanism of punishment is a result of motivation.

In this model we explain cooperative behavior as a result of a moral motivation. This motivation is, however, not exogenous but can be influenced by the agents.

[5]See Axelrod (1984) and the literature on the Folk Theorem, summarized for example in Fudenberg and Tirole (1991).

[6]See Hirshleifer (1999) for a classification of different types of cooperation problems.

[7]See Alexander (1987).

Hence, cooperative behavior results from an altruistic motive *ex post*, but this altruistic motive is created by egoistic agents *ex ante* given the technological constraints of the economy. The moral disposition, that means the ability to create a mutual concern, is assumed to be exogenously given, whereas the extend to which morality becomes important for actual decisions is seen as an endogenous reaction to the environment. In this sense, morality or altruism are substitutes for institutions in the narrower sense like property rights with associated law enforcement. This hypothesis is supported by the following observations. (a) There exists an interest of individuals in topics like morality, justice etc. If these acts of communication are more than entertainment or reassurance, their aim has to be to influence behavior. (b) One can observe the attempt of societies to create common values by the education system, the churches, or the media. These activities consume time and other resources and the question arises why a society would be interested in making these investments.

Morality becomes important in situations of conflict. In a conflict-free society morality is irrelevant. In a conflict the distribution of goods and resources is determined by power and not by considerations about overall efficiency, and it induces the waste of resources invested to appropriate and to defend against appropriation.[8] The paper builds on this literature to shows that voluntary investments in a morality can reduce the extend of such a conflict by creating an intrinsic motivation to avoid stealing and betrayal.

A related approach that is based on conflict theory is the paper by Grossman and Kim (2001) which analyzes the relationship between morality and economic performance. In this paper the exogenous moral disposition explains the structure of the equilibrium and characterizes conditions under which morally motivated agents are rewarded for their disposition. This paper's concept of a reward of morality is different because the agents voluntarily invest to create a mutual concern. One interesting consequence of this approach is that it allows to understand a mutual concern as the result of underlying technological factors.

We derive three main results. First we show that initially egoistic agents have in fact an incentive to create a mutual concern if the enforcement of property rights is

[8]See Grossman (2001), Hirshleifer (1995), and Skaperdas (1992) among others.

costly. The amount of resources invested in the creation of morality depends on the technology of production as well as on the technology of appropriation and conflict. Hence, second we show that the investments in morality are increasing in the productivity of the economy, which implies that *ceteris paribus* more productive regions have a stronger mutual concern. Third, morality becomes most important if the abilities to appropriate and to defend are relatively equal. Perhaps most surprisingly it can be shown that the investments in a mutual concern are decreasing if defense is becoming increasingly difficult. Hence, morality is a good substitute for other means of enforcement only for specific conflict technologies.

The paper proceeds as follows. In Section 2 we present the model and the basic equilibrium conditions. In Section 2.1 we characterize the equilibrium for a given morality, and in Section 2.2 we characterize the equilibrium with endogenous morality. Section 3 concludes.

2 The model

Consider an economy populated by two (representative small) economic agents $i = 1, 2$.[9] Markets in this economy are not initially ensured and costlessly implemented but are created by the agents.

We assume that each agent is endowed with $L \in [1, \infty)$ units of time that he can use for five different activities. Agent i uses a quantity l_i to produce a (private) good, y_i. Total production and the production function are denoted by $y_i = F(l_i) = l_1^x$, $x \in (0, 1)$. The parameter x measures the natural productivity of the economy.

Agent i can also try to appropriate the good produced by the other agent. We denote by a_i the amount of time devoted to appropriation. Knowing that the other agent will try to appropriate the good, agent i can also defend against appropriation. d_i is the amount of time devoted to the defense of his possession. In addition the agents can choose to be idle and consume f_i units of time as leisure. The last activity

[9]The assumption that both agents are small means that they treat prices as parameters.

is the investment of m_i units of time in the establishment of a morality. We denote by $\delta_i = \{l_i, d_i, a_i, f_i, m_i\}$, $i = 1, 2$, and $\delta = \{\delta_1, \delta_2\}$ the vector of decision variables.

A crucial part of this model is the way the concept of a moral disposition is operationalized. We define a *moral disposition* as the mental ability to develop an awareness and concern about the effects of one's behavior upon other individuals. Accordingly, an individual is *moral* if it uses its moral disposition to internalize the effects of its behavior upon other individuals to a certain degree and to change its behavior accordingly.[10] We denote by b the parameter that measures the degree of internalization, or morality.

In order to define the individual objective functions we distinguish between the direct utility function and the individualistic welfare function of the individual.[11] The direct utility function is defined on consumption, c_i^1, c_i^2, and leisure, f_i, and is homothetic,

$$u_i = u(c_i^1, c_i^2, f_i) = c_i^1 \cdot c_i^2 \cdot f_i, \quad i = 1, 2. \tag{1}$$

Given the parameter of mutual concern, b, the individualistic welfare function is defined on both individuals' direct utility functions and assumed to be utilitarian,

$$W_i = W_i(c_i^1, c_i^2, f_i, c_j^1, c_j^2, f_j, b) = c_i^1 \cdot c_i^2 \cdot f_i + b(c_j^1 \cdot c_j^2 \cdot f_j), \quad i = 1, 2. \tag{2}$$

This formulation of mutual concern is structually equivalent to Hamilton's Rule, where b plays to role of the relatedness parameter r defined in the introduction. The difference between both models is that r depends on genetic relatedness and is exogenous from the point of view of the individual, whereas b can be influenced by the individuals. Hirshleifer (1985) gives an alternative interpretation of the parameter b. In his conception, b measures the balance of power between an altruistic and a selfish personality. To be more specific we assume that $b = b(m_1, m_2) = m_1 \cdot m_2$.[12]

[10] This definition of a moral disposition fits the purpose of this paper to develop a theory of market transactions supported by morality and not outside enforcement. A more general concern for the well being of other individuals may exist but is not relevant for this paper. Please note that the distinction between different types of moral concerns are irrelevant in a two-agent economy.

[11] See Harsanyi (1955) and Roemer (1996) for a discussion of the concept of individualistic welfare functions.

[12] A simple extension of this model in the form $b = b(m_1, m_2, r)$ would allow to incorporate both,

The iso-elastic formulation reflects two plausible properties of the creation of morality. First, the creation of a moral concern can only be successful if both agents participate, $m_1 > 0, m_2 > 0$. Second, the creation of mutual concern is easier if both agents invest a more equal amount of time, $b(\lambda m_1, (1-\lambda)m_2) > b(m_1, m_2)$, $\lambda \in (0,1)$, $m_1 \neq m_2$. These properties reflect the idea that the creation of a mutual concern is a cooperative act in the sense that the agents are more successful if they behave similarly. The investment in this communication process is voluntary for both agents. The creation of a mutual concern is in their own interest, otherwise they will not act. In Hamilton's theory our model implies that the agents can deviate from the degree of mutual concern predetermined by genetic relatedness. Using Hirshleifer's interpretation our model assumes that the balance between the selfish and the altruistic personality can be changed in favor of the altruistic type if the individuals increase the amount of resources devoted to the creation of morality.

We denote by $\pi_i \in [0,1]$ the fraction of good i that can be successfully defended against the appropriation of agent j by agent i.[13] We assume that this fraction is determined by the Tullock contest-success function,

$$\pi_i = \pi_i(d_i, a_j, \theta) = \frac{1}{1 + \theta a_j / d_i}, \quad \theta \in (0, \infty). \tag{3}$$

θ is a parameter that characterizes the technology of conflict. Larger values of θ *ceteris paribus* increase the fraction of the good that is appropriated,

$$\frac{\partial \pi_i}{\partial \theta} = -\frac{a_j/d_i}{(1 + \theta a_j/d_i)^2} < 0.$$

At the same time, a larger value of θ may de- or increase the marginal effectiveness

Hamilton's approach and the idea that the individuals can influence their mutual concern to some extend. In such a model, r would create a 'level' effect.

[13]This formulation of the problem implies that the ability to appropriate or defend is influenced by the initial production of the agents. In contrast to the case where the production has no influence on the position of the individuals in the contest (which is called *common-pool problem* by Grossman (2001)) it makes sense to have both, appropriation and trade with such a model setup.

of appropriative as well as defensive activities, depending on the sign of $(\theta a_j - d_i)$,

$$\begin{aligned} \frac{\partial^2 \pi_i}{\partial d_i \partial \theta} &= \frac{a_j(\theta a_j - d_i)}{(d_i + \theta a_j)^3}, \\ \frac{\partial^2 \pi_i}{\partial a_j \partial \theta} &= \frac{d_i(d_i - \theta a_j)}{(d_i + \theta a_j)^3}. \end{aligned}$$

The presence of the possibility to appropriate goods does not rule out the possibility of regular exchange. Typically, both activities, trade and theft, will coexist. It is question of the opportunity costs whether it is more attractive to buy or to steal. Our model captures this observation by the following sequential structure of the game:

Stage 1: At stage 1 the agents allocate (non-cooperatively) their time budget to communication, productive, defensive, and appropriative activities as well as to leisure. This defines a primary distribution of the private goods, $E = \{E_1^1(\delta), E_1^2(\delta), E_2^1(\delta), E_2^2(\delta)\}$ with $E_1^1 = \pi_1 y_1$, $E_2^1 = (1 - \pi_1) y_1$, $E_2^2 = \pi_2 y_2$, $E_1^2 = (1 - \pi_2) y_2$, where subscripts on variables denote the agent and superscripts denote the good.

Stage 2: At stage 2 the agents can trade their endowment under conditions of perfect competition (i.e. both agents take prices as parameters).

The game is solved by backward induction.

Stage 2: At stage 2 both agents take as given the outcome of the contest at stage 1. Denote by $c = \{c_1^1, c_1^2, c_2^1, c_2^2\}$ the final consumption levels and by p_2 the relative price of good 2 in terms of good 1. The agents maximize their individualistic welfare function by the choice of c_i^1, c_i^2:[14]

$$\max_{c_i^1, c_i^2} W_i(c_i^1, c_i^2, f_i, c_j^1, c_j^2, f_j, b) \quad s.t. \quad c_i^1 + p_2 c_i^2 = E_i^1 + p_2 E_i^2. \tag{4}$$

[14] Technically it makes no difference whether we assume that the agents maximize their individualistic welfare functions or their utility functions, the Marshallian demand functions are the same for the same structure of markets. However, from a normative point of view it makes a difference whether the individuals evaluate their market choices with their welfare or their utility functions. If the consumption values of the other agent are an argument of the target function their exists an interdependence that would call for individualized prices for the individual levels of consumption in order to internalize them (Arrow 1977). All relevant interdependencies are internalized by the

We denote by λ_i the Lagrangian multiplier associated with the budget constraint. The first order conditions,

$$\frac{\partial W_i/\partial c_i^2}{\partial W_i/\partial c_i^1} = p_2, \quad i = 1, 2, \tag{5}$$

show that the marginal rates of substitution coincide, there is efficiency in exchange irrespective of the instability of property rights due to possibility of appropriation. Markets insure the efficiency of exchange even in the presence of appropriation. Appropriation restricts the extend of the market, but the relevant transactions to ensure *ex-post* efficiency still take place.

The solution to this problem gives rise to Marshallian demand functions

$$c_1^1(\delta) = \frac{E_1^1(\delta) + p_2 E_1^2(\delta)}{2}, \tag{6}$$

$$c_1^2(\delta) = \frac{E_1^1(\delta) + p_2 E_1^2(\delta)}{2p_2}, \tag{7}$$

$$c_2^1(\delta) = \frac{E_2^1(\delta) + p_2 E_2^2(\delta)}{2}, \tag{8}$$

$$c_2^2(\delta) = \frac{E_2^1(\delta) + p_2 E_2^2(\delta)}{2p_2}. \tag{9}$$

A market equilibrium at stage 2 is a price $p_2(\delta)$ such that $c_1^j(\delta, p_2) + c_2^j(\delta, p_2) = E_1^j(\delta) + E_2^j(\delta)$, $j = 1, 2$. It is given by

$$p_2(\delta) = \frac{E_1^1(\delta) + E_2^1(\delta)}{E_1^2(\delta) + E_2^2(\delta)}. \tag{10}$$

Inserting this price into the Marshallian demand functions gives rise to equilibrium demand functions, $c_i^j(\delta)$, $i, j = 1, 2$. Inserting these functions into the individualistic welfare functions and using the definitions of $E_1^1(\delta)$, $E_1^2(\delta)$, $E_2^1(\delta)$, $E_2^2(\delta)$ yields

market structure in the case that the evaluation takes place by the use of the utility functions. This problem of interpretation is relevant for every model building on the first theorem of welfare economics because it neglects the empirical fact that individuals are mutually concerned. The purpose of this paper is to demonstrate that morality can emerge in equilibrium and how it is influenced by the underlying technological structure. We therefore do not provide any attempt to solve this problem here.

indirect welfare functions

$$V_1(\delta) = \frac{l_1^x l_2^x f_1(d_1 d_2 + a_1\theta(2d_1 + a_2\theta))^2}{4(d_2 + a_1\theta)^2(d_1 + a_2\theta)^2} + b\frac{l_1^x l_2^x f_2(d_1 d_2 + a_2\theta(2d_2 + a_1\theta))^2}{4(d_2 + a_1\theta)^2(d_1 + a_2\theta)^2},$$

$$V_2(\delta) = \frac{l_1^x l_2^x f_2(d_1 d_2 + a_2\theta(2d_2 + a_1\theta))^2}{4(d_2 + a_1\theta)^2(d_1 + a_2\theta)^2} + b\frac{l_1^x l_2^x f_1(d_1 d_2 + a_1\theta(2d_1 + a_2\theta))^2}{4(d_2 + a_1\theta)^2(d_1 + a_2\theta)^2}.$$

Stage 1: At stage 1 the agents maximize their indirect welfare functions $V_i(\delta)$ by the choice of δ_i, $i = 1, 2$.

A Nash equilibrium of the game at stage 1 is a vector $\bar{\delta}_i^N = \{a_i^N, d_i^N, l_i^N, f_i^N, m_i^N\}$, $i = 1, 2$, such that $\bar{\delta}_i^N \in \arg\max_{\bar{\delta}_i} V_i(\bar{\delta}_i, \bar{\delta}_j^N, t)$ s.t. $l_1 + d_i + a_i + f_i + m_i = 1$, $i = 1, 2$, $j \neq i$. Due to the symmetry of the problem we know that a symmetric equilibrium with $\bar{\delta}_1^N = \bar{\delta}_2^N$ exists. After some simplifications the first-order conditions for agent 1 can be written as

$$\frac{\partial V_1}{\partial a_1} = -(\lambda_1 - b\lambda_2)\frac{\partial \pi_2}{\partial a_1}F_2 - \frac{\partial u_1}{\partial f_1} + \lambda_1\frac{\partial p_2}{\partial a_1}(E_1^2 - c_1^2), \tag{11}$$

$$\frac{\partial V_1}{\partial d_1} = (\lambda_1 - b\lambda_2)\frac{\partial \pi_1}{\partial d_1}F_1 - \frac{\partial u_1}{\partial f_1} + \lambda_1\frac{\partial p_2}{\partial d_1}(E_1^2 - c_1^2), \tag{12}$$

$$\frac{\partial V_1}{\partial l_1} = (\lambda_1\pi_1 - b\lambda_2(1 - \pi_1))\frac{\partial F_1}{\partial l_1} - \frac{\partial u_1}{\partial f_1} + \lambda_1\frac{\partial p_2}{\partial l_1}(E_1^2 - c_1^2), \tag{13}$$

$$\frac{\partial V_1}{\partial m_1} = b\frac{\partial u_2}{\partial c_2^1}\left(\frac{dE_2^1}{dm_1} + p_2\frac{dE_2^2}{dm_1}\right) + \frac{\partial b}{\partial m_1}u_2 - \frac{\partial u_1}{\partial f_1} + \lambda_1\frac{\partial p_2}{\partial m_1}(E_1^2 - c_1^2). \tag{14}$$

Analogous conditions can be derived for agent 2. Before we give a precise interpretation of the first-order conditions we start with an interpretation of the last terms in (11) - (14), $\lambda_1\partial p_2/\partial q_1(E_1^2 - c_1^2)$, $q_1 = a_1, d_1, l_1, m_1$. These terms measure the general-equilibrium effect of a change in time investment on the equilibrium price of the economy. This strategic effect is similar to the strategic incentives resulting from a hold-up problem (Grossman and Hart 1986) and occurs because we have restricted attention to an economy with only two agents. In such an economy every agent anticipates that his allocation of time has an influence of the future equilibrium price. We will interpret the results assuming that this effect is zero in the following in order to concentrate on the remaining effects.

In the following we will first interprete the structure of the equilibrium for a given level of b in order to be able to determine the general influence of a mutual

concern on the equilibrium (constrained Nash equilibrium). In a second step we will then analyze the time investments of the agents in morality (full Nash equilibrium).

2.1 The constrained Nash equilibrium

What is the interpretation of (11) - (13)? The partial effects of a change in a_1 and d_1 are similar and given in (11) and (12). An increase in a_1 increases the fraction of the good produced by agent 2 that is possessed by agent 1. The marginal utility of the increased possession is given by $(\lambda_1 - b\lambda_2)$. It is composed out of the marginal utility from the consumption of good 1 for agent 1, λ_1, and the morally internalized effect on the marginal utility of agent 2, $b\lambda_2$. The utility costs of an increase in appropriation are given by the second term, the marginal utility of leisure. In an interior solution appropriative investments are chosen such that both effects are equal. The net effect, $(\lambda_1 - b\lambda_2)$, shows the economic role of a moral disposition in this model: appropriation and defense create a negative-sum game. A mutual concern of b partially internalizes the adverse effect for agent 2 of an increase in appropriation by agent 1. This shows how a moral disposition can help to overcome the inefficiency of markets with costly enforcement of property rights. For the benchmark of a transaction-costs free enforcement of property rights, $\theta = 0$, $\partial\pi_2/\partial a_1 = 0$, appropriation is impossible. This is the benchmark case of general-equilibrium theory. In this benchmark no agent invests any resources in appropriation because the marginal utility is equal to zero, whereas the marginal costs are positive. A similar result can be achieved with a perfect mutual concern, $b = 1$.

The symmetry of our model gives $\lambda_1 = \lambda_2$, which implies the following result:

> **Result 1:** (a) In a symmetric Nash equilibrium, no agent invests in appropriation and defense if and only if $b = 1$. (b) For a given strategy of agent 2 (1), an increase in b decreases appropriative and defensive investments of agent 1 (2).

Proof: (a) The if-part is obvious because the first term in (11) and (12) cancel if $\lambda_1 = \lambda_2$. For the only-if part note that the Tullock function has the property

$\partial\pi_1(0,0,\theta)/\partial d_1 \to \infty \; \forall\theta \neq 0$. This implies that for $b < 1$ the first term dominates the second and third terms at $d_1 = d_2 = a_1 = a_2 = 0$. (b) An increase in b decreases the marginal value of appropriation and defense. At the same time the marginal utility of leisure remains constant, which implies the result. qed

Finally we interpret the incentives to produce, which are determined by (13). Every agent equates the weighted marginal productivity to the marginal utility of leisure. The weight $\lambda_1(\pi_1 - b(1-\pi_1))$ attached to the marginal productivity is equal to the weighted average between both agents' marginal utility of consumption. The weight depends (a) on the fraction of theft and (b) the mutual concern for the other agent. If $b = 0$ (mutually unconcerned agents) the condition reduces to $\lambda_1\pi_1$, if $b = 1$ the condition reduces to λ_1. In the latter case the agent invests until the marginal rate of substitution between consumption and leisure is equal to the marginal rate of transformation between consumption and leisure. This is the first-best condition known from general-equilibrium theory. It follows that there exist incentives for underinvestment in production as long as the mutual concern is imperfect. This result stems from the fact that each agent has to bear the full marginal costs of production, whereas he internalizes only part of the marginal gains because a fraction of his production is stolen. This distortion of marginal conditions compared to the first best creates a second source of inefficiency in this model (in addition to the direct resource costs of the contest).

We can now determine the equilibrium values for the constrained Nash equilibrium by a simultaneous solution of (11)-(13) and the analogous conditions for agent 2. We use $b = m_1 \cdot m_2$ and the symmetry of the equilibrium:

$$a_i(m_i,\theta,x,L) = \frac{2(L - m_i - Lm_i^2 + m_i^3)\theta}{1 + x + m_i^2 x + 2\theta(3m_i^2(x-2) + x) + \theta^2(1 + x + m_i^2 x)}, \quad (15)$$

$$d_i(m_i,\theta,x,L) = \frac{2(L - m_i - Lm_i^2 + m_i^3)\theta}{1 + x + m_i^2 x + 2\theta(3m_i^2(x-2) + x) + \theta^2(1 + x + m_i^2 x)}, \quad (16)$$

$$l_i(m_i,\theta,x,L) = \frac{-(-L + m_i - Lm_i^2 + m_i^3)(1+\theta)^2 x}{1 + x + m_i^2 x + 2\theta(3m_i^2(x-2) + x) + \theta^2(1 + x + m_i^2 x)}, \quad (17)$$

and $f_i = L - l_i - a_i - d_i - m_i$. The solution allows us to derive more detailed results about the relationship between mutual concern, productivity, and the structure of the equilibrium:

Result 2: Defensive and appropriative investments are decreasing in productivity and morality,

$$\frac{\partial a_i^N}{\partial x} < 0, \frac{\partial d_i^N}{\partial x} < 0,$$
$$\frac{\partial a_i^N}{\partial m_i} < 0, \frac{\partial d_i^N}{\partial m_i} < 0.$$

The proof of the result can be found in the appendix. What is the interpretation of this result? Figure 1 illustrates the finding for the example of $L = 10$, $\theta = 1/2$, and $m_i = 1/2$. The decreasing function measures the investment in appropriation, a_i, whereas the increasing function measures labor input. An increase in productivity

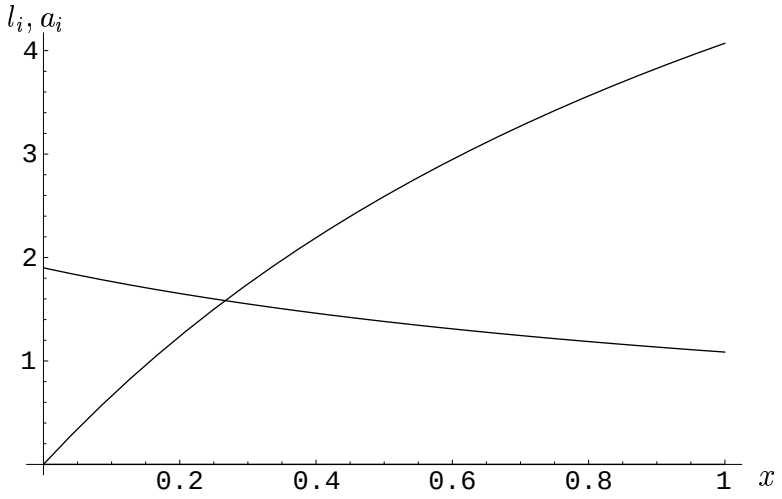

Figure 1: Appropriative and productive investments as functions of productivity.

x increases the opportunity costs of appropriative investments. At the optimum the marginal increase in utility from productive and appropriative/defensive investments is equal. A marginal increase in labor productivity *ceteris paribus* increases the return from labor, which explains part (a) of the result. This finding has important implications for the structure of conflicts that are empirically observed. Our model explains why for a given level of b economies with a smaller natural productivity tend to have more intensive conflicts. This implies that economies with high productivity

do not only profit from the better production possibilities but also from the induced reduction in conflict.

Part (b) of the result seems to confirm the hypothesis that it is the right type of morality that favors economic prosperity that can first be found in Weber's seminal study on the relationship between protestant ethics and economic performance. According to Weber a religion or ethic that supports capitalist behavior is a key factor for economic prosperity, and he identifies three crucial elements of such an ethic: the willingness to work hard, the willingness to accumulate capital, and the development of a character and an idea of charity that implies the voluntary compliance with contractual obligations.[15] The voluntary compliance with contractual obligations allows the widespread establishment of market transactions that could otherwise not be supported by the legal system.[16]

Figure 2 illustrates the finding for the example of $L = 10$, $\theta = 1/2$, and $x = 1/2$. The decreasing function measures the investments in appropriation, whereas the increasing function measures labor input. An increase in concern for the well being of the other agent internalizes to a greater degree the externality that appropriative and defensive activities have on the other agent. In this sense morality is a substitute for explicit institutions. The purpose of institutions is to shape incentives in a way that every agent internalizes (marginally) the total effect of his actions upon other agent. Explicit institutions like the the concept of property rights plus the enforcement agency ensuring that contracts are not violated influence behavior by influencing

[15]See Weber (1988), pp. 207. Weber provides evidence that the affiliation to the baptist and methodist churches was a signal of economic trustworthiness in the United States of the 19th century.

[16]There are numerous theories of moral behavior in the European and American history that support the view that compliance with mutual obligations is a key element of moral behavior. Immanuel Kant is a famous advocate of *pflichethik*, an ethic based on duty, which has as a key element the idea that individuals are bound to their promises irrespective of the consequences. The idea of a social contract as the foundation of a just society rests on the idea of the obligational power of voluntary promises, see for example Hobbes (1994), and for a review contractarian models Kolm (1996). From the perspective of our model these theories can be seen as attempts to overcome the cooperation problem of society resulting from enforcement problems of property rights by creating the 'right' individual motivation.

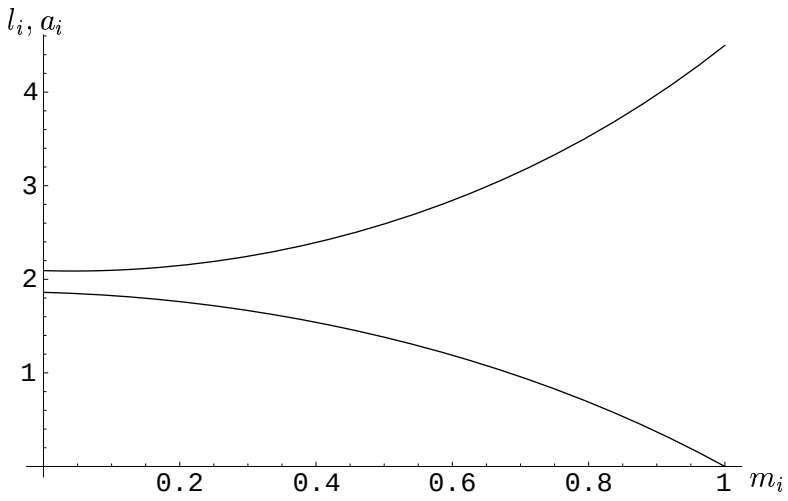

Figure 2: Appropriative and productive investments as functions of mutual concern.

the restrictions of the agents. Morality can be seen as an implicit or internalized constraint with the same purpose and potentially the same effects as institutions in the narrower sense of the word. If the society is successful in the establishment of a mutual concern it is successful in the reduction of wasteful appropriative and defensive activities.

2.2 The full Nash equilibrium

We now turn to the analysis of the full Nash equilibrium in order to see if (a) the agents actually invest in the creation of a mutual concern and (b) how these investments depend on the underlying technology. This endogenization challenges the causality between morality and prosperity. If the assumption is shared that morality is an endogenous institutional response of the agents to mitigate the potential for conflict in allocation problems, it is not the mutual concern that drives economic prosperity but natural productivity that drives the mutual concern, and by this channel prosperity.

In order to determine the full Nash equilibrium the simultaneous solution of (11)-

(14) and the analogous conditions for agent 2 has to be determined. It is straightforward to show that the solution is given by

$$m_i^N = \frac{L(1+\theta)^2}{2(\theta-1)^2+(\theta+1)^2x} - \frac{\sqrt{(L^2(\theta+1)^4-4((\theta-1)^2+(\theta+1)^2x)(1+x+\theta(6+\theta+(\theta+2)x)))}}{2(\theta-1)^2+(\theta+1)^2x}. \quad (18)$$

An evaluation of (18) yields:

> **Result 3:** There are positive investments in the creation of morality in equilibrium. The equilibrium investment in the creation of morality are increasing in productivity, $\partial m_i/\partial x > 0$.

Proof: The derivative of (18) with respect to x is

$$\frac{\partial m_i^N}{\partial x} = \frac{(1-\theta)^2 \frac{-L(1+\theta)^2+(-16(\theta-1)^2\theta+L^2(1+\theta)^4-16\theta(1+\theta)^2x}{\sqrt{L^2(1+\theta)^4-4((\theta-1)^2+(1+\theta)^2x)(1+x+\theta(6+\theta+(2+\theta)x))}}}{2((\theta-1)^2+(1+\theta)^2x)^2}. \quad (19)$$

In order to determine the sign of this derivative assume that there exists a point such that $\partial m_i/\partial x = 0$. If this point exists, it is characterized by

$$x = \frac{-1+2\theta-\theta^2}{(1+\theta)^2}.$$

For $\theta \in [0,1)$ it follows that $x < 0$, which implies that (19) is either strictly positive or strictly negative for all meaningful parameter values $x \in [0,1]$, $L \in [1,\infty)$, and $\theta \in [0,1)$. It is therefore sufficient to characterize the derivative for one specification of the parameter values in order to determine the sign for the whole range of parameters. For example for $\theta = 0.5$, $L = 10$, and $x = 0.5$, we get $\partial m_i^N/\partial x = 0.109114$. qed

Result 3 is represented by Figure 3. The Figure is drawn for $L = 10$ and $\theta = 1/2$. The intuition for the positive relationship between natural productivity and morality is that a rise in productivity increases the production potential in a conflict-free economy. This implies that for a given level of morality an rise in productivity increases the absolute loss induced by the contest. An increase in morality is a means to partly counterbalance this effect.

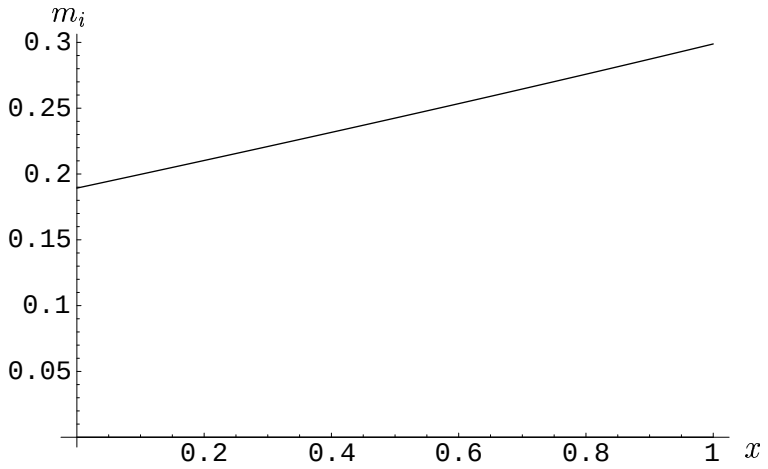

Figure 3: The influence of productivity on the mutual concern.

Next it is shown how the investment in a mutual concern is influenced by the technology of conflict, which is summarized by θ.[17] A larger value of θ *ceteris paribus* implies a larger marginal and absolute effectiveness of appropriation, which corresponds to situations where the defense of property is increasingly difficult.

> **Result 4:** The investments in a moral disposition are maximal if both agents are equally strong in the contest, $\theta = 1$. It is increasing before and decreasing thereafter. The investments in a moral concern are the same for $\theta = 0$ and $\theta \to \infty$. A sufficient condition for this solution to exist is $L \geq 4$.

[17] See Hirshleifer 1985 for a discussion of the technology of conflict.

Proof: The derivative of (18) with respect to t is

$$\frac{\partial m_i^N}{\partial \theta} =$$
$$- \frac{2(\theta-1)^2 - L^2(1+\theta)^2 + 4(1+x)((\theta-1)^2+(1+\theta)^2 x)}{((\theta-1)^2+(1+\theta)^2 x)^2 \sqrt{L^2(1+\theta)^4 - 4((\theta-1)^2+(1+\theta)^2 x)(1+x+t(6+\theta+(2+\theta)x))}}$$
$$- \frac{L\sqrt{L^2(1+\theta)^4 - 4((\theta-1)^2+(1+\theta)^2 x(6+\theta+(2+\theta)x)}}{((\theta-1)^2+(1+\theta)^2 x)^2 \sqrt{L^2(1+\theta)^4 - 4((\theta-1)^2+(1+\theta)^2 x)(1+x+t(6+\theta+(2+\theta)x))}}. \quad (20)$$

Setting this equation equal to zero yields $\theta = 1$. Checking the limit behavior yields

$$\lim_{\theta \to 0} m_i^N = \lim_{\theta \to \infty} m_i^N = \frac{L - \sqrt{L^2 - 4(1+x)^2}}{2(1+x)}.$$

An inspection of both expressions shows that they are fulfilled irrespective of x for $L \geq 4$. If $L < 4$ there are parameter values x such that the square root becomes negative. qed

Result 4 is represented by Figure 4. The Figure is drawn for $L = 10$ and $x = 1/2$.

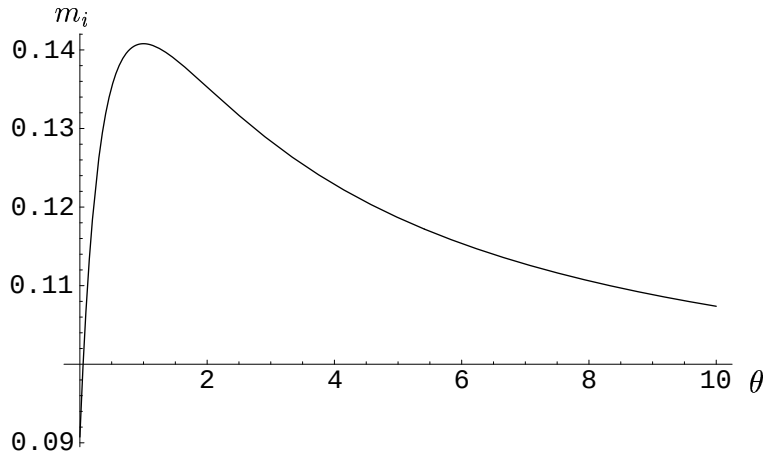

Figure 4: The influence of decisiveness on morality.

Result 4 may seem surprising. One might have argued that morality is becoming more important if the defense of property is becoming more difficult. The result,

however, suggests that morality can play a role for 'intermediate' values of θ only. The intuition for this finding is as follows. If appropriation is impossible ($\theta = 0$) there is no need for a mutual concern in this framework. Adam Smith's intuition that mutually unconcerned individuals cooperate to their mutual advantage by the invisible hand of the market turns out to be correct. If appropriation becomes possible but the defender has still a comparative advantage in defense ($\theta \in (0,1)$), an increasing amount of resources has to be invested in order to avoid appropriation. This implies that a mutual concern becomes important. However, if defense is becoming more and more difficult it is very easy to appropriate. This high effectiveness of appropriation induces the appropriator to invest less in the contest, which in turn implies that the investments in defense are reduced. Hence, for high values of the decisiveness parameter the total resource costs of the contest go down. In the limit it is impossible to defend against appropriation and the appropriator gets the goods for an infinitesimal investment in appropriation. In this situation the resource costs of creating a morality that is sufficient to prevent such an behavior are prohibitive.[18] This finding shows that if the creation of morality requires the investments of resources, in general the agents decide to invest in the creation of an imperfect morality. Morality is a means to solve a cooperation problem that is used until the marginal cost of morality (measured by the loss of production) is equal to the marginal gain from morality (resulting from a reduced intensity of conflict).

Another interesting implication of Result 4 is that the agents have an incentive to invest in morality even if appropriation is impossible. We know that the investments are minimized but weakly positive at that point. For the example in Figure 4 we get $m_i|_{\theta=0} \approx 0.075$, which implies $b|_{\theta=0} \approx 0.0056$. This 'minimum morality' is a result of the fact that an increase in mutual concern increases the utility not only because of a reduction in appropriation and defense, but also because the agents 'care' about the well being of the other agent directly. The fact that this 'nosieness' of the agents is minimized with stable property rights shows that the main incentive to invest in a mutual concern results from the imperfection of the market. This result, however, shows also that a mutual concern and perfect property rights are

[18] This finding results from the paradox of power first characterized by Hirshleifer (1991).

no perfect substitutes. If the agents start investing in a mutual concern they cannot avoid to become interested in the well-being of the other agent.

3 Conclusions

In this paper I have developed a theory of the rational development of morality among economic agents. There are some implications of the result that should further be stressed. First, initially egoistic individuals have a rational incentive to create a mutual concern in order to reduce the resource costs of conflict. Moral behavior in this sense is a means for the effective organization of a society in a world where the enforcement of property rights incurs a cost: morality helps the invisible hand of the market to establish efficiency. In a transaction-costs free society a moral disposition would be obsolete to achieve the efficiency of markets, it is in this world where Adam Smith's insight turns out to be most powerful. At the other end of the spectrum, with perfect concern for the other agent, both agents voluntarily refrain from stealing because the negative effect on the other agent is perfectly internalized. None of the agents tries to steal because no one wants to steal. This implies that the agents do not have to invest in defense and, more importantly, every transaction in this economy is carried out by the use of markets where the mutual obligations are voluntarily met without the need for private enforcement or public enforcement.

Second, the strength of the morality depends on both, the technology of production and the technology of appropriation and defense. Economies with a large natural productivity are more likely to develop a morality that favors the creation of a comprehensive system of markets. This implies that a morality supporting markets is perhaps not the explanation of economic prosperity but itself the consequence of a highly productive economy. Our results show that the natural productivity of a region causes both, the development of a favorable morality and economic prosperity. This finding is in accordance with Weber (1988) who mentions the fact that the protestant revolution was first observed in the economically most highly developed regions of Europe.[19] This finding is able to explain why for example British colo-

[19]See Weber (1988), pp.19.

nialism was successful in the United States but not in Africa. If morality were the explanatory factor there should be no systematic differences. This view is supported by the current findings in growth theory. Engerman and Sokoloff (1997), McArthur and Sachs (2001), Sachs (2001) among others strongly suggest that ecological and climatic characteristics together with key geographical factors like access to ports are the best explanatory variable for economic development. Our findings add a new element to this discussion in claiming that geography, and thereby productivity, might also be responsible for the extend of conflict and the development of a morality supporting markets because every agent can rely on the voluntary fulfillment of contractual obligations.

However, it is not only the productive technology that is important for the creation of morality but also the conflict technology. Our results support the view that morality becomes important if both agents are relatively equal in their abilities to defend and appropriate. If defense is much easier than appropriation there is no need for a mutual concern because markets work well anyway. If on the other hand appropripation is very easy the resource costs of the contest are relatively low. Nevertheless, the distortion of the incentives to produce are high, but the resource costs of creating the strength of mutual concern necessary to fill the gap are prohibitive.

Appendix: Proof of Result 2

First note that $d_1^N = a_1^N$. We can therefore restrict attention to a_1^N. The derivative of a_1^N with respect to x is

$$\frac{\partial a_1^N}{\partial x} = \frac{2(L-m_1)(m_1^2-1)(m_1^2+1)\theta(1+\theta)^2}{(1+6\theta-4m_1^2\theta+\theta^2+(1+m_1^2)(1+\theta)^2x)^2}.$$

The denominator is always positive. The numerator can only be equal to 0 if $L = m_1$. However, $m_1 \leq L$ from the time constraint. Denote by $m_1 = L - k$, $k \in [0, m_1]$ the equilibrium time investment. The numerator can then be written as $((k-L)^2 - 1)(1+(k-L)^2)\theta(1+\theta)^2 2k$. The sign of this expression depends on the sign of the first term, $(k-L)^2 - 1$. This term is negative if $k > L+1$. $k > L+1$ implies $m_1 < -1$, a contradiction.

The derivative of a_1^N with respect to m_1 must have the same denominator as the derivative of a_i^N with respect to x. hence, it is also positive. The numerator is equal to

$$\begin{aligned}\Psi =: \quad & - \quad 2\theta(1+4m_1^4\theta+2Lm_1(1+\theta)^2+\theta(6+\theta)-m_1^2(3+\theta(14+3\theta))) \\ & - \quad 2(1+4Lm_1-m_1^2(4+m_1^2))\theta(1+\theta)^2x. \end{aligned} \tag{A.1}$$

Note that (A.1) is continuously differentiable in all parameters. We proof the result in two steps. We first show that the derivative is negative at $m_i = 0$ and $m_i = 1$. Then we show that there exists no point such that the derivative is equal to zero. Both properties and continuity imply that the derivative is negative for all $m_1 \in [0,1]$.

(1) The lower and upper bound for m_1 is 0 (no mutual concern, $b = 0$) and 1 (complete mutual concern, $b = 1$) respectively. For these values (A.1) simplifies as follows:

$$\begin{aligned}\left.\frac{\partial a_1^N}{\partial m_1}\right|_{m_1=0} \leq 0 \quad &\Leftrightarrow \quad -2\theta(1+x+\theta(6+\theta+(2+\theta)x)) \leq 0, \\ \left.\frac{\partial a_1^N}{\partial m_1}\right|_{m_1=1} \leq 0 \quad &\Leftrightarrow \quad -4(L-1)\theta(1+\theta)^2(1+2x) \leq 0.\end{aligned}$$

Both conditions are always fulfilled.

(2) Next we show that there exists no interior value $m_1 \in (0,1)$ such that $\partial a_1/\partial m_1 = 0$. First of all, (A.1) is monotonous in x. The term multiplied by x,

$$\Phi = 1+4Lm_1-m_1^2(4+m_1^2),$$

is minimized for $L = 1$. At $L = 1$ it follows immediately that $\Phi \geq 0$ for all $m_1 \in [0,1]$. This implies that the derivative $\partial a_i/\partial m_i$ is largest at $x = 0$. We therefore analyze the structure of Ψ at $x = 0$. (A.1) reduces to

$$\Psi|_{x=0} = -2\theta(1+4m_1^4\theta+2Lm_1(1+\theta)^2+\theta(6+\theta)-m_1^2(3+\theta(14+3\theta))).$$

Assume an interior solution exists, it has to fulfill

$$L = -\frac{1+4m_1^4\theta+\theta(6+\theta)-m_1^2(3+\theta(14+3\theta))}{2m_1(1+\theta)^2}.$$

The resource constraint implies $(L - m_i) \geq 0$, which yields

$$\begin{aligned}(L - m_1) &= -\frac{(m_1^2 - 1)(-1 - \theta(6 - 4m_1^2 + \theta))}{2m_1(1+\theta)^2} \geq 0 \\ &\Leftrightarrow -(m_1^2 - 1)(-1 - \theta(6 - 4m_1^2 + \theta)) \geq 0. \qquad \text{(A.2)}\end{aligned}$$

It is straightforward to show that this inequality cannot be fulfilled for $m_1 \in (0, 1)$ and $\theta \in [0, \infty)$. This implies that every point where

$$\left.\frac{\partial a_1}{\partial m_1}\right|_{x=0} = 0$$

must violate the resource constraint. We already know that $\partial a_1 / \partial m_1 < 0$ is fulfilled at $m_1 = 0$ and $m_1 = 1$, which completes the proof. qed

4 References

Alexander, R., 1987. The Biology of Moral Systems. Aldine de Gruyber, New York.

Arrow, K.J., 1977. The Organization of Economic Activity: Issues Pertinent to the Choice of Market Versus Nonmarket Allocation, in: R.H. Havemann, J. Margolis (Hrsg.): Public Expenditure and Policy Analysis, Chicago.

Axelrod, R., 1984. The Evolution of Cooperation. Basic Books, New York.

Bush, W.C., Mayer, L.S., 1974. Some Implications of Anarchy for the Distribution of Property. Journal of Economic Theory 8 (4), 401-412.

Danielson, P., 1996. The Rights of Chickens: Rational Foundations for Libertarianism. In: Sanders, J.T., Narveson, J. (Eds.): For and Against the State. Rowman and Littlefield, London.

Engerman, S., Sokoloff, K., 1997. Factor Endowments, Institutions, and Differential paths of Growth among the New World Economies. In: Haber, S. (ed.): How Latin America Fell Behind. Stanford University Press.

Fudenberg, D., Tirole, J., 1991. Game Theory. MIT-Press, Cambridge.

Gauthier, D., 1986. Morals by Agreement. Oxford University Press, Oxford.

Grossman, H., 2001. The Creation of Effective Property Rights. American Economic Review 91, 347-352.

Grossman, H., Kim, M. 2001. Is a Moral Disposition Rewarded? Journal of Banking and Finance, forthcoming.

Hamilton, W.D., 1964. The Genetical Evolution of Social Behavior, 1. Journal of Theoretical Biology 7, 1-17.

Harsanyi, J.C., 1955. Cardinal Welfare, Individualistic Ethic, and Interpersonal Comparison of Utility. Journal of Political Economy 69, 309-321.

Hirshleifer, J., 1985. The Expanding Domain of Economics. American Economic Review 75, 53-68.

Hirshleifer, J., 1991. The Paradox of Power. Economics and Politics 3, 177-200.

Hirshleifer, J., 1995. Anarchy and its Breakdown. Journal of Political Economy 103 (1), 26-52.

Hirshleifer, J., 1999. There are many Evolutionary Pathways to Cooperation. Journal of Bioeconomics 1, 73-93.

Hobbes, T. 1994 [1668]. Leviathan. Indianapolis.

Kagel, J.H., Roth, A.E., 1995. The Handbook of Experimental Economics. Princeton, New Jersey.

Kolm, S.C., 1996. Modern Theories of Justice. Cambridge.

McArthur, J.W., Sachs, J.D., 2001. Institutions and Geography: Comment on Acemoglu, Johnson and Robinson (2000). NBER Working Paper 8114.

Narveson, C., 1988. The Libertarian Idea. Temple University Press, Philadelphia.

Roemer, J., 1996. Theories of Distributive Justice. Cambridge.

Sachs, J.D., 2001. Tropical Underdevelopment. NBER Working paper 8119.

Skaperdas, S., 1992. Cooperation, Conflict, and Power in the Absence of Property Rights. American Economic Review 82 (4), 720-739.

Smith, A., 1976 [1759]. The Wealth of Nations. Random House, New York

Weber, M., 1988 [1920]. Gesammelte Aufsätze zur Religionssoziologie I. J.C.B. Mohr UTB (Paul Siebeck), Tübingen.